ELSEVIER

Nuclear Instruments and Methods in Physics Research A 389 (1997) 210–213

NUCLEAR INSTRUMENTS & METHODS IN PHYSICS RESEARCH Section A

Offline analysis of HEP events by "dynamic perceptron" neural network

A.L. Perrone [a,b,c,*], G. Basti [b,c], R. Messi [a,b], E. Pasqualucci [a,b], L. Paoluzi [a,b]

[a] *Department of Physics, University of Rome "Tor Vergata" Via della Ricerca Scientifica 1, I-00133, Rome, Italy*
[b] *INFN, National Institute for Nuclear Physics, Section of Rome 2 Via della Ricerca Scientifica 1, I-00133, Rome, Italy*
[c] *Pontifical Lateran University Piazza S. Giovanni in Laterano, 4, I-00184, Rome, Italy*

Abstract

In this paper we start from a critical analysis of the fundamental problems of the parallel calculus in linear structures and of their extension to the partial solutions obtained with non-linear architectures. Then, we present shortly a new dynamic architecture able to solve the limitations of the previous architectures through an automatic re-definition of the topology. This architecture is applied to real-time recognition of particle tracks in high-energy accelerators.

PACS: 84.35; 13.85; 97.60.J
Keywords: Offline analysis; Pattern recognition; Neural network; Neutron form factor

1. The theory

The design of an effectively parallel computation architecture is actually the main problem to be solved, in order to perform efficient complex pattern recognition tasks. This limitation has its best mathematical formulation in the classical theorems discussed in Minsky and Papert *Perceptrons* book [1]. Essentially, these theorems demonstrate the mathematical impossibility of designing *truly parallel* architectures for solving very simple recognition problems (e.g., "parity" problem or "xor" problem). In fact, to solve these problems with a *linear* perceptron, it is necessary to have at least one computation unit (e.g., a "neuron" or a set of "neurons" in the hidden layer of a neural architecture) with the whole dimensionality of the input as its computation domain.

Generally, the researchers followed two different approaches for dealing with this type of problems:

– The use of *classical AI* strategies with objects already defined by the programmer (i.e. object oriented languages) [2].

– The use of a *connectionist approach*, i.e. a neural architecture (e.g., "backpropagation") with two main differences with respect to the classical linear perceptron: (a) the use of a *non-linear* transfer function for each neuron in order to generate higher-order input correlations [3]; and (b) the *total connectivity* among the neurons on the different layers.

Notice that in a connectionist architecture the limitations of the linear preceptron are not solved in principle because the total connectivity implies that each neuron of the hidden layer "reads" the whole input. In this way, the computation *is not truly parallel*. The only improvement is that this architecture can solve non-linearly separable problems such as the "xor problem". However, no theorem exists for granting the convergence of the learning procedure in a finite time. In the applications a long time for the convergence of the learning procedure is thus generally required.

In this paper we propose another strategy to solve the problem of parallel computation. This strategy is based on a *dynamic definition of the net topology* that shows its effectiveness for problems of particle track recognition in high-energy physics [4,5]. In this way, we can preserve the linear architecture of the net like in the classic perceptron, but with a partial and dynamic connectivity so to overcome the intrinsic limitations of the geometric perceptron. Namely, the computation is truly parallel

* Correspondence address: Department of Physics, University of Rome "Tor Vergata" Via della Ricerca Scientifica 1, I-00133, Rome, Italy. Tel.: +39 6 4742529; fax: +39 6 2023507/ +39 6 4742529; e-mail: perrone@roma2.infn.it.

PII S0168-9002(97)00118-6

because of the partial connectivity, but the net topology is always the optimal one because of its dynamic redefinition on each single input pattern. For these properties, we call this new architecture *dynamic perceptron*.

To synthesize the main differences among: (1) the classic geometric perceptron, (2) the back-propagation algorithm and (3) the dynamic perceptron, it is sufficient to consider the following three schemes:

(i) Geometric perceptron: *Fixed and partial* input correlations for the single neuron are defined according to the following (see Fig. 1):

$$\Psi(X) = \left[\sum_{i=1}^{N-1} x_i x_{i+1} > \theta\right], \tag{1}$$

where $[x > \theta] = 1$ if $x > \theta$, and $[x > \theta] = 0$ if $x \leq \theta$.

(ii) Back propagation: *Fixed and total* input correlations for the single neuron are defined according to the following (see Fig. 2):

$$o_k = f\left(\sum_{j=1}^{N_k} z_{kj} \cdot f\left(\sum_{i=1}^{N_j} w_{ji} x_i\right)\right), \tag{2}$$

where z_{kj} is the weight between the jth hidden unit and the kth output unit; w_{ji} is the weight between the ith input unit and the jth hidden unit; x_i is the vector forcing directly the input units, N_j is the number of inputs for each hidden unit; N_k is the number of inputs for each output unit: f is the sigmoid function defined by $f(x) = 1/(1 + e^{-x})$. The correlation is *total* because of the following simple Taylor expansion, where all the products among all the input neurons $x_i, x_{i_1}, x_{i_2}, \ldots$,; $i, i_1, i_2, \ldots = 1, \ldots, N$ occur:

$$f\left(\sum_{i=1}^{N} w_{ki} x_i\right) \simeq c_1\left(\sum_i w_{ki} x_i\right) + \frac{c_2}{2}\left(\sum_i \sum_{i_1} w_{ki} w_{ki_1} x_i x_{i_1}\right) + \frac{c_3}{6}\left(\sum_i \sum_{i_1} \sum_{i_2} \cdots\right) + \cdots . \tag{3}$$

(iii) Dynamic perceptron: *Dynamic and partial* correlations, bounded by an upper limit, are defined for the single neuron. By "dynamic" connectivity we mean a time-varying connectivity of the neurons of the net. Namely, the connectivity changes for *each* input pattern so to give the *optimal topology* to "read" it, according to the following (see Fig. 3):

$$\Psi^D(X) = \left[\sum_{i=1}^{m} p_i(X) > \theta\right], \tag{4}$$

where the input supports $S_{p_i}(X)$ of the computation unit p_i (i.e., the part of the whole input X used as input of each p_i) are recursively constructed for each input pattern X,

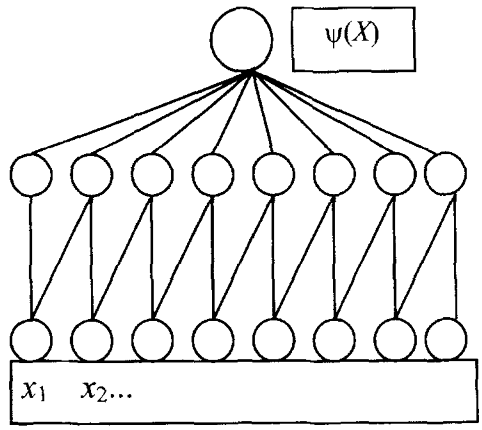

Fig. 1. Order two correlations in the geometric perceptron.

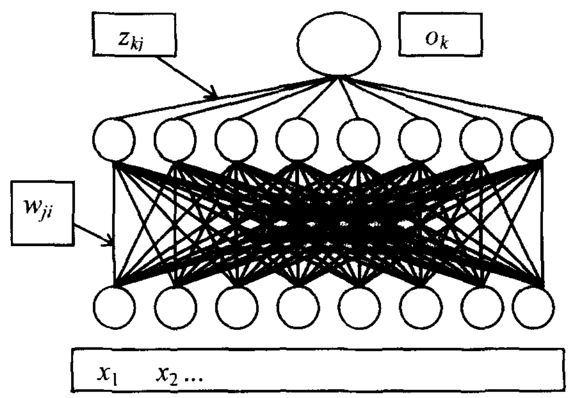

Fig. 2. Total correlation between hidden and input layers for the back-propagation.

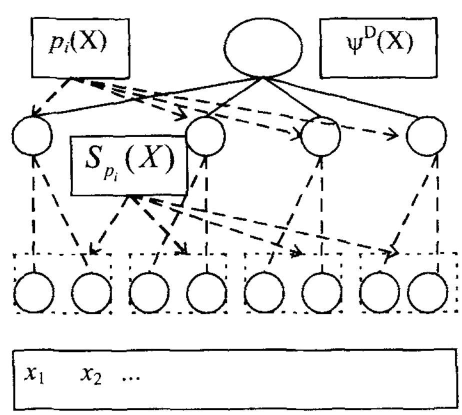

Fig. 3. Dynamic and partial correlations for the dynamic perceptron. The connectivity is free to change in each box delimiting a upper limit of order two to the correlations.

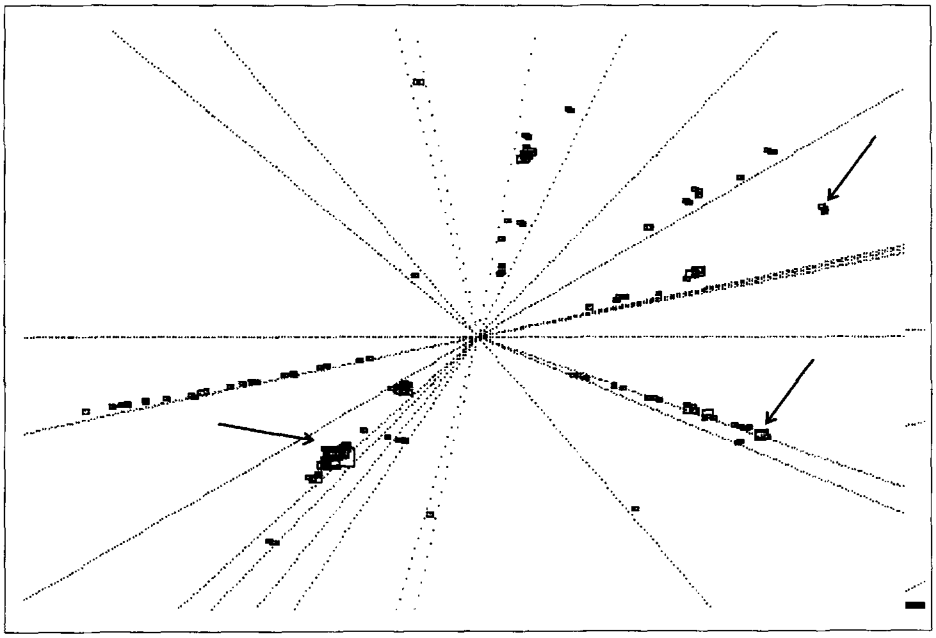

Fig. 4. Antineutron star after the dynamic redefinition of the supports to determine the connected or uniform parts of the track. The little squares indicated by the arrows show some of these supports redefined on connected areas of the input pattern.

so to determine connected or uniform parts of the single input (see Fig. 4).

2. An application

The "FENICE" detector, installed at the upgraded storage ring ADONE from '89 to '93, is a non-magnetic detector and consists of a complex array of plastic scintillators, iron sheets and streamer tubes. It is optimized to detect the process $e^+e^- \to n\bar{n}$ and consequently to measure the neutron electromagnetic form factors. A concrete shield covered by an active veto system is added in order to reduce the cosmic rays background. Due to the very small cross section, the expected rate is of the order of $1\text{–}10^6$ triggers at the luminosities of $\approx 10^{29}\,\mathrm{cm}^{-2}\,\mathrm{s}^{-1}$.

Three types of data were submitted to the dynamic perceptron in order to test its efficiency in recognizing different classes of events: *raw data*; *Monte Carlo events*; *real events*.

Raw data: They consist of topological information obtained only from the streamer tubes (ADC, TDC, ... give other topological information, but with a more difficult utilization) and concerning events surviving a first reduction filter aimed to reduce the presence of cosmic rays. The recognition task consists essentially here in distinguishing collinear and non-collinear charged events – occurring in the proximity of the interaction zone – from the background events of $e^{\pm}$–gas interactions, in order to extract events in which a "star" indicating an annihilation of an antineutron is present.

Monte Carlo events: They simulate the reaction $e^+e^- \to n\bar{n}$ at different energies from 1920 to 3100 MeV.

Real events: They consist in samples of real $e^+e^- \to n\bar{n}$ annihilation events obtained through human scanning. In these samples, events are accepted only if they are judged as satisfying the criteria for a "star" recognition in the event by at least 3 over 6 human scanners. The results obtained by the dynamic perceptron for all these tasks of complex pattern recognition are summarized in Table 1.

Notice that the class of Monto Carlo rejected events ($\approx 15\%$ of the total) is composed as follows: 10% quasi-collinear events; 5% events with too few points. Finally, the efficiency of the new architecture is enhanced by the reduction of calculation time. The above results is obtained in *real time by software* (15 Hz of recognition on 300×300 input points on a 4 Mips machine).

Table 1
Summary of the results obtained by the dynamic perceptron

Events	Accepted (%)	Rejected (%)
Monte Carlo (1.000 events)	≈ 85%	≈ 15%
Real data (≈ 400 events at ≈ 3.1 GeV)	≈ 90%	≈ 10%
Raw data (≈ 100.000 events)	40–50%	50–60%

3. Conclusions: The advantages of the method

The dynamic perceptron displays the following advantages with respect to *any other* neural network architecture designed till now:

– It uses only *the topological information* (no ADC, flight time, etc.).

– It allows *a complete parallelization* of the algorithm.

– It makes *independent* the net structure from the apparatus dimension (i.e., the net rearranges automatically itself on any change of the apparatus dimensions).

– It allows the treatment of *any input dimension in real time by software* (15 Hz of recognition on 300×300 input points on a 4 Mips machine).

– It allows an *easy hardware implementation* for the exceptional architectural simplicity of the net (actually the hardware implementation is in development).

References

[1] M.L. Minsky and S. Papert, Perceptrons, Expanded Edition (MIT Press, Cambridge, MA, 1982) p. 56.
[2] M.L. Minsky, The Society of Minds (MIT Press, Cambridge, MA, 1982).
[3] D.E. Rumelhart, G.E. Hinton and R.J. Williams, in: Parallel Distributed Processing Vol. 1 eds. D.E. Rumelhart and J.L. McClelland (MIT Press, Cambridge, MA, 1982) p. 318.
[4] G. Basti, A.L. Perrone, P. Castiglione and R. Messi, in: W. Ruck eds., Applications of Artificial Neural Networks, V, eds. S.K. Rogers and D.W. Ruck, SPIE-Proc. Series, 2243 (SPIE-The Int. Soc. for Optical Engineering, Bellingham, WA, 1994) p. 540.
[5] A.L. Perrone, Lecture Notes in Computer Science 888 (1995) 9.